

Historic, Archive Document

Do not assume content reflects current scientific knowledge, policies, or practices.

aTX360
U6N385
October
19-20,
1988

CONTINUING SURVEY OF FOOD INTAKE BY INDIVIDUALS - 1989 (CSFII)

- SAMPLE DESIGN -

prepared for

**HNIS
UNITED STATES DEPARTMENT OF AGRICULTURE**

prepared by

NATIONAL ANALYSTS
A Division of Booz•Allen & Hamilton Inc.

OCTOBER 19-20, 1988

THE DESIGN OF THE CSFII 1989 SAMPLE MUST BE SCIENTIFICALLY SOUND AND VERSATILE. THE SAMPLE NEEDS TO:

- Provide projectable, valid information about food consumption patterns in American households
- Yield data about seasonal variations in dietary patterns
- Net data that can be linked to past and future survey efforts by USDA and other federal agencies
 - CSFII 1985 and 1986
 - NFCS 1987
 - CSFII 1990, 1991, 1992
 - CSFII Elderly
 - HANES
 - Others

CSFII SAMPLE OVERVIEW: THERE ARE NINE STEPS IN THE CSFII SAMPLING EFFORT; EACH REQUIRES CAREFUL PLANNING AND CONTROL

- Identify appropriate nationally-representative sample frame
- Draw two independent samples of area segments for complete CSFII activities 1989-1992 and assign segments to years
- Classify according to income levels in low-income survey
- Assign area segments to season
- Select sample housing units for contact
- Screen and interview at all eligible households
- Select certainty sample of family members for intake interviewing
- Select and interview respondents to DHK questionnaire
- Weight resultant data to national total

APPROPRIATE SAMPLE FRAME

NATIONAL ANALYSTS' MASTER SAMPLE FRAME IS APPROPRIATE BECAUSE IT IS...

- A multistage area probability sample of the conterminous United States
- Stratified by geography and urbanization
- Based upon 1980 Census data, updated to 1985
- Four independent, interlocking samples of up to 240 primary sampling units (PSUs)

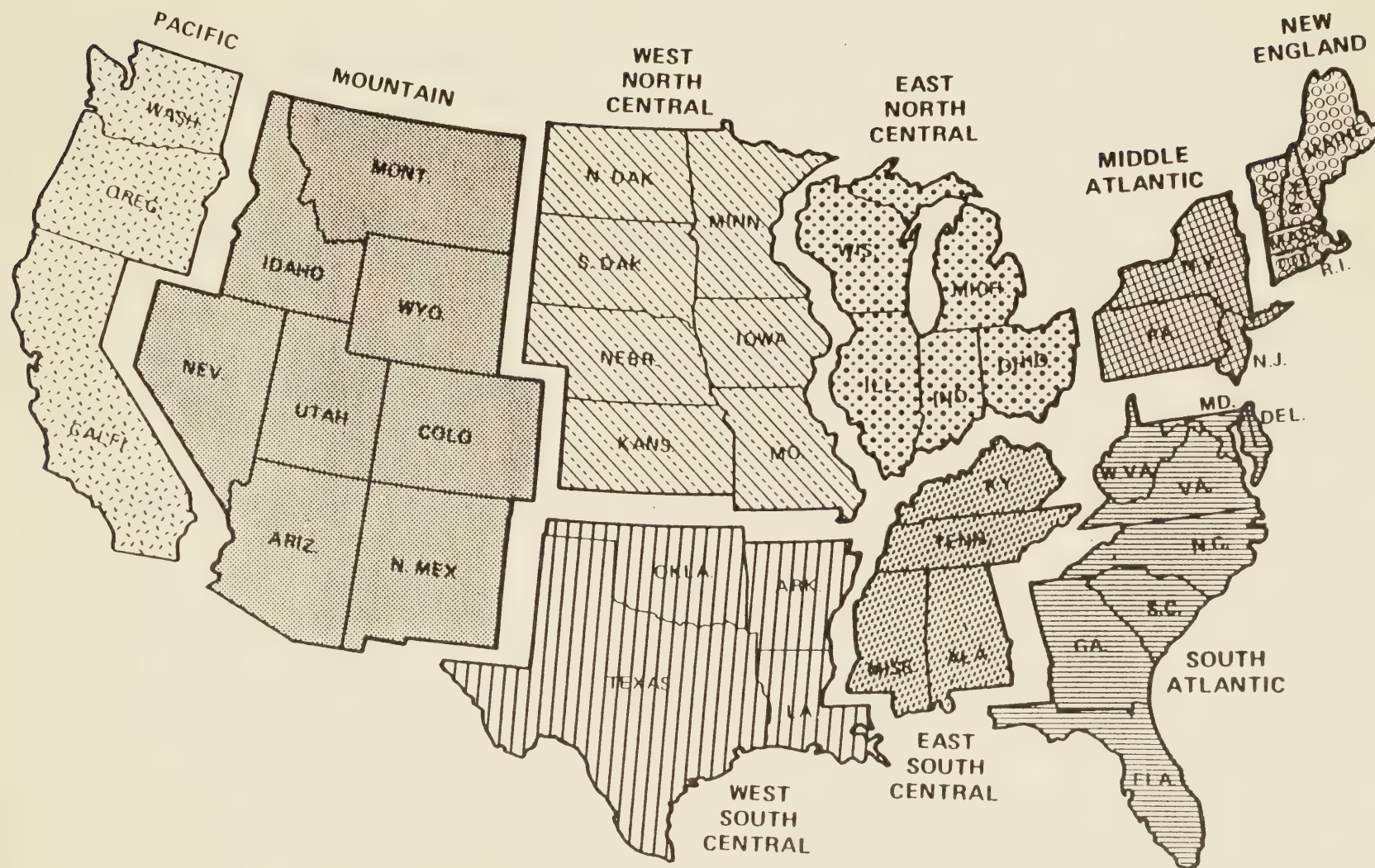
APPROPRIATE SAMPLE FRAME

STRATIFICATION OF NATIONAL ANALYSTS' MASTER SAMPLE...

- Premised upon:
 - Census divisions (9)
 - Urbanization (3)
- This 27-cell matrix provides framework for creating 60 nonoverlapping strata
- All housing units and people within the conterminous United States are located in one of the 60 strata

APPROPRIATE
SAMPLE FRAME

CENSUS DIVISION STRATA



APPROPRIATE SAMPLE FRAME

URBANIZATION STRATA ARE...

- Metropolitan areas: Central cities (427 in total)
- Metropolitan areas: Central city county minus central city (696 counties or county equivalents in suburbia)
- Nonmetropolitan areas (area outside SMAS)

APPROPRIATE SAMPLE FRAME

WITHIN URBANIZATION AND DIVISION STRATIFICATION 60 STRATA FORMED...

- Strata formed on the basis of:
 - Population size
 - Geographic proximity
 - Other homogenizing factors (e.g., economics)
- Each of the 60 strata represents an average of about 4,000,000 persons



1



APPROPRIATE SAMPLE FRAME

STRATIFICATION OF THE UNITED STATES BY DIVISION
AND URBANIZATION INTO STRATA

<u>Division</u>	<u>Metro Central City</u>	<u>Metro Suburban</u>	<u>Nonmetro</u>	<u>Total</u>
New England	1	1	1	3
Middle Atlantic	3	5	1	9
E North Central	3	6	2	11
W North Central	1	1	2	4
E South Central	1	1	2	4
South Atlantic	2	5	3	10
W South Central	2	3	2	7
Mountain	1	1	1	3
Pacific	3	5	1	9
Total	17	28	15	60

APPROPRIATE SAMPLE FRAME

EACH STRATUM SIZE BASED UPON ESTIMATED 1985 POPULATION...

- 1985 population estimated by:
 - Measuring ratio of growth for two five-year periods --
1970/1975 and 1975/1980
 - Projecting five-year growth to 1985 from known 1980 Census
data
- Growth rates estimated for every central city, suburban county and
nonmetro county in conterminous United States

APPROPRIATE SAMPLE FRAME

SELECTION OF THE MASTER SAMPLE PRIMARY SAMPLING UNITS (PSUs) IS MADE WITH PROBABILITY PROPORTIONAL TO SIZE (PPS)...

- Size estimates based on projected 1985 population for the PSU and stratum
- Four replicate PSUs selected with replacement for each of the 60 strata

APPROPRIATE
SAMPLE FRAME

NATIONAL ANALYSTS NEW ENGLAND
CENTRAL CITY STRATIFICATION

- PSU SELECTION -

<u>CENTRAL CITY</u>	<u>STATE</u>	<u>ESTIMATED POPULATION</u>	<u>CUMULATIVE POPULATION</u>
Bangor	ME	31,031	31,031
Lewiston	ME	39,922	70,953
Auburn	ME	22,948	93,901
Portland	ME	63,546	157,447
Manchester	NH	100,336	257,783
Nashua	NH	75,638	333,421
Portsmouth	NH	27,971	361,392
Dover	NH	23,380	384,772
Rochester	NH	23,756	408,528
Burlington	VT	38,333	446,861
Boston	MA	504,956	951,817
Brockton	MA	94,526	1,046,343
Fall River	MA	86,134	1,132,477
Fitchburg	MA	40,273	1,172,750
Leominster	MA	33,644	1,206,394
Worcester	MA	152,859	1,359,253
Lawrence	MA	59,497	1,418,750
Haverhill	MA	49,749	1,468,499
Lowell	MA	93,390	1,561,889
New Bedford	MA	96,851	1,658,740
Pittsfield	MA	49,256	1,707,996
Springfield	MA	138,220	1,846,216
Chicopee	MA	52,319	1,898,535
Holyoke	MA	42,923	1,941,458
Providence	RI	146,611	2,088,069
Warwick	RI	88,375	2,176,444
Pawtucket	RI	70,347	2,246,791

APPROPRIATE SAMPLE FRAME

NATIONAL ANALYSTS NEW ENGLAND
CENTRAL CITY STRATIFICATION

- PSU SELECTION -
 (Continued)

<u>CENTRAL CITY</u>	<u>STATE</u>	<u>ESTIMATED POPULATION</u>	<u>CUMULATIVE POPULATION</u>
Bridgeport	CT	142,095	2,388,886
Bristol	CT	56,288	2,445,174
Danbury	CT	67,300	2,512,474
Hartford	CT	134,430	2,646,904
Meriden	CT	56,568	2,703,472
New Britain	CT	69,414	2,772,886
New Haven	CT	125,319	2,898,205
West Haven	CT	53,367	2,951,572
New London	CT	27,339	2,978,911
Norwich	CT	35,464	3,014,375
Stamford	CT	99,802	3,114,177
Waterbury	CT	99,700	3,213,877
Norwalk	CT	78,914	3,292,791

TOTAL STRATUM SIZE 3,292,791

APPROPRIATE SAMPLE FRAME

WE WILL USE 120 PSUs FOR CSFII 1989-1992

- Consistent with prior CSFII and NFCS projects
- Provides USDA with two sample replicates for variance estimation

APPROPRIATE SAMPLE FRAME

SHOULD ALASKA AND HAWAII BE INCLUDED?

APPROPRIATE SAMPLE FRAME

INCLUSION OF ALASKA AND HAWAII ADDRESSES SOME ISSUES AND PRESENTS SOME PROBLEMS

PROS

- Allows projectability to total population of the U.S.
 - Alaska: 2 strata
 - Hawaii: 3 strata -- central city, suburbia, nonmetro
- Makes data base more comparable to HANES
- ???

CONS

- Need to draw and staff new PSUs
- Must support survey with additional nutrient data bases; adds complexity
- ???

DRAW TWO SETS OF SEGMENTS

**ALL AREA SEGMENTS FOR THE BASIC AND LOW-INCOME CSFII 1989-1992
SURVEYS WILL BE IDENTIFIED THIS YEAR TO AVOID DUPLICATION AND OVERLAP**

- A minimum of 250 area segments in the basic sample will be fielded each year
- A minimum of 218 area segments in the low-income sample will be fielded each year
- Cross-year duplication of segments will be controlled. Cross-sample overlap may occur, but duplication will be avoided at the sample housing unit selection stage

DRAW TWO SETS OF SEGMENTS

THE STEPS IN THE SELECTION OF AREA SEGMENTS ARE...

- Determine the total number of area segments

- 1,000 basic segs for CSFII 1989-1992
- 870 low-income segs for CSFII 1989-1992

(identify 2,000 and
subsample 870)

- Determine the distribution of area segments by PSU

- $n_i = k_i N/P_i$ where n_i = number of segments

k_i = overall segment sampling rate

N = 1985 estimated population of PSU

P_i = probability of selection of PSU

DRAW TWO SETS OF SEGMENTS

STEPS...

- Select all sample area segments in a PSU
 - Array all housing units systematically
 - Determine sampling interval in PSU
 - Draw random number starting point (number between 1 and the sample interval)
 - Select appropriate number of housing units
 - Identify housing units selected and area segments containing them
- Assign sample area segments to one of four CSFII survey years

(Continued)

DRAW TWO SETS OF SEGMENTS

(Continued)

STEPS...

WORCESTER, MA (III)

KINGS, NY (112)

WORCESTER, MA (III)					KINGS, NY (112)				
<u>Year</u> <u>Code</u>	<u>Selection</u> <u>Random</u> <u>Number</u>	<u>Tract</u>	<u>Block</u>	<u>HUs</u>	<u>Year</u> <u>Code</u>	<u>Selection</u> <u>Random</u> <u>Number</u>	<u>Tract</u>	<u>Block</u>	<u>HUs</u>
8	503	7301	301-307	100	7	36997	50	201	211
1	9305	7305	412-414	97	8	95720	110	102	141
2	18107	7309.02	304-306	78	1	154443	165	402	292
3	26909	7313	401-403	95	2	213166	220	102	222
4	35711	7318	402-403	85	3	271889	267	401	235
5	44513	7322.03	403-405	97	4	330612	317.01	301	301
6	53315	7327	205-206	135	5	389335	360.02	202	415
					6	448058	416	104	176

- Year codes 1 and 5
- 2 and 6
- 3 and 7
- 4 and 8

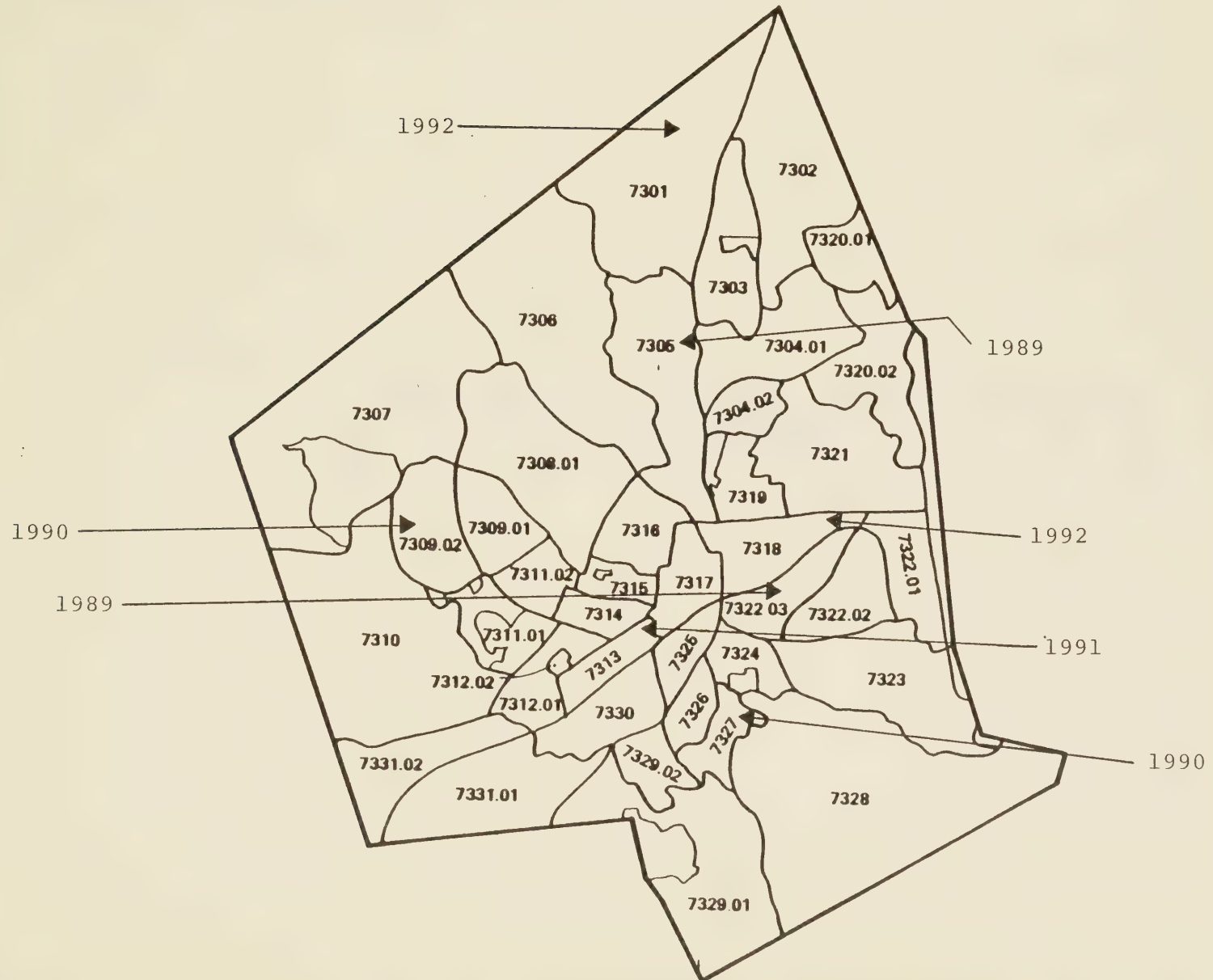
CSFII 1989

CSFII 1990

CSFII 1991

CSFII 1991²

WORCESTER, MASSACHUSETTS AREA SEGMENTS



DRAW TWO SETS OF SEGMENTS

**WHEN AND HOW SHOULD ELDERLY AREA SEGMENTS
BE IDENTIFIED?**

DRAW TWO SETS OF SEGMENTS

ELDERLY SEGMENTS...

- Could be separate and identified later -- possible overlap
- Could be separate and identified now -- no overlap, but added costs now
- Could be part of low-income sample segments CSFII 1989-1991 or 1992

CLASSIFY LOW-INCOME SEGMENTS

TO COMPLETE WORK IN THE LOW-INCOME SAMPLE EFFICIENTLY, WE WILL UNDERSAMPLE AREAS WITH LITTLE KNOWN POVERTY AND OVERSAMPLE AREAS OF KNOWN POVERTY

- We will identify 1,850 area segments for the low-income sample
- All these segments will be classified by NPD Census data into one of three poverty strata
 - High 25% or more poverty
 - Medium 10% to 24% poverty
 - Low Less than 10% poverty
- Area segments to be fielded will be subsampled differentially by poverty status

	<u>Segment Selection Rate</u>	<u>Number of Segments Fielded</u>
- High	1.00	370
- Medium	0.40	348
- Low	0.25	<u>152</u>
		870

SEASONAL SEGMENT ASSIGNMENT

BECAUSE THE NUMBER OF INTERVIEWS TO BE COMPLETED IN EACH AREA SEGMENT IS SMALL BUT CONSISTENT WITH PRIOR SURVEY TARGETS, WE PROPOSED THAT HALF OF THE SEGMENTS BE FIELDDED EACH QUARTER

- Target number of basic interviews: 6 per segment
 1.5 per quarter

- Target number of low-income interviews: 3+ per segment
 less than 1 per quarter

- To control the flow of work throughout the year efficiently:

	SPRING	SUMMER	FALL	WINTER
BASIC	<div style="border: 1px solid black; padding: 5px;"> 125A Segments 3 Int/Qt 375 Interviews </div>	<div style="border: 1px solid black; padding: 5px;"> 125B Segments 3 Int/Qt 375 Interviews </div>	<div style="border: 1px solid black; padding: 5px;"> 125A Segments 3 Int/Qt 375 Interviews </div>	<div style="border: 1px solid black; padding: 5px;"> 125B Segments 3 Int/Qt 375 Interviews </div>
LOW- INCOME	<div style="border: 1px solid black; padding: 5px;"> 110A Segments Nearly 2 Int/Qt 188 Interviews </div>	<div style="border: 1px solid black; padding: 5px;"> 110B Segments Nearly 2 Int/Qt 188 Interviews </div>	<div style="border: 1px solid black; padding: 5px;"> 110A Segments Nearly 2 Int/Qt 188 Interviews </div>	<div style="border: 1px solid black; padding: 5px;"> 110B Segments Nearly 2 Int/Qt 188 Interviews </div>

SEASONAL SEGMENT ASSIGNMENT

SHOULD WE FIELD HALF THE SEGMENTS EACH QUARTER?

SELECT SAMPLE HOUSING UNITS

THE NUMBER OF SAMPLE HOUSING UNITS TO BE SELECTED IS A FUNCTION OF...

- Actual number of housing units in 1989 as enumerated on listing form
- Expected occupancy rate (91%)
- Targeted number of completed household interviews (basic = 1,500; low-income = 750)
- Expected eligibility rate (basic = 100%; low-income = 25%)
- Estimated completion rate (75%)

SELECT SAMPLE HOUSING UNITS

ONCE THE TOTAL NUMBER OF SAMPLE HOUSING UNITS IS DETERMINED, THE
NUMBER TO BE CONTACTED IN A SEGMENT WILL BE IDENTIFIED

- Sample housing units will be selected with known probability
 - Will be self-weighting in the basic sample

- We will share with USDA
 - The probability of selection of each sample housing unit for both the basic and low-income samples

SCREEN AND INTERVIEW ELIGIBLES

ALL HOUSEHOLDS WILL BE CONTACTED AND SCREENED TO DETERMINE ELIGIBILITY

- All occupied housing units in the basic sample are eligible for interview
- Only income/size qualifying households are eligible for interview in the low-income sample
- Special efforts will be made to complete interviews among eligibles
 - Nonresponse will be controlled at division and urbanization levels
- Data will be captured for noninterviewed households

SCREEN AND INTERVIEW ELIGIBLES

**WHAT ARE THE INCOME/SIZE GUIDELINES FOR
THE LOW-INCOME SAMPLE?**

CERTAINTY SAMPLE FOR INTAKE

ALL HOUSEHOLD MEMBERS WILL COMPLETE UP TO THREE DAYS OF INTAKE RECORDINGS IN THE BASIC AND LOW-INCOME SAMPLES. MISSING AND INCOMPLETE DATA ISSUES NEED TO BE RESOLVED. WHAT IF...

- One or more members do not complete the intake?
- Some days are missing for some individuals?
- Days are not consecutive? Are different for different members?
- Days are partial; complete as far as they go?

SELECT DHK RESPONDENT

ONE DHK TELEPHONE INTERVIEW WILL BE ATTEMPTED WITH A MEMBER OF EVERY
CSFII-INTERVIEWED HOUSEHOLD

- Who should be the respondent in the DHK survey
 - Main meal planner/preparer?
 - Other random adult?

WEIGHTING

THE RESULTANT HOUSEHOLD AND INTAKE DATA WILL BE WEIGHTED TO NATIONAL POPULATION FIGURES

- Weights will be developed to:
 - Allow each year's data to stand alone
 - Permit merging basic and low-income samples
 - Combine across years for rolling average
- Our weighting model is premised upon occupied sample housing units
 - Build-up to sample housing units will be by segment by quarter
 - Household interviews will be adjusted to housing unit division totals
 - Intake interviews will be adjusted to age and sex national totals

WEIGHTING

**WHAT ARE THE OTHER WEIGHTING PROBLEMS THAT
HAVE ARISEN IN THE PAST AND SHOULD BE
ADDRESSED IN THE PRESENT CSFII?**

